

The Fingertip Effects of Computer-based Assessment in Education

By Hong Lin and Francis Dwyer

According to the Committee on the Foundations of Educational Assessment, traditional educational assessment does a reasonable job of measuring knowledge of basic facts, procedures and proficiency of an area of the curriculum. However, the traditional approach fails to capture the breadth and richness of knowledge and cognition (Pellegrino, Chudowsky, & Glaser, 2001). Such a concern arises because traditional assessment practices generally focus on assessing whether a student has acquired the content knowledge, but they often fail in assessing the learning process and higher-order thinking skills (Baek, 1994; Bahr & Bahr, 1997). Dede (2003) concludes that the current practices of educational assessment are “based on mandating performance without providing appropriate resources, then using a ‘drive by’ summative test to determine achievement” (p. 6).

At a time when traditional assessment is under increasing scrutiny and criticism, the nation is placing greater expectations on the potential role of the computer in educational assessment. It is anticipated that the appropriate use of computer technology would help enhance assessment at multiple levels of practice by incorporating ongoing and multiple assessment strategies into the learning process. Given this possibility, it is timely to review current computer-based assessment practices in educational settings. Furthermore, a review of some emerging assessment tools that incorporate interactive multimedia can also deepen our understanding of the role that computer technology plays in assessment.

Technology use in assessment and fingertip effects

Fingertip effects of computer technology

Computer technology has significantly changed the curriculum and teachers’ instructional practices. It has also changed the way students construct and demonstrate their knowledge and skills. These changes “in turn are stimulating people to rethink what is assessed, how that information is obtained, and how it is fed back into the educational process in a productive and timely way” (Pellegrino, Chudowsky, & Glaser, 2001, p. 272).

Perkins (1985), a pioneer thinker who viewed computers as learning tools, pointed out that computer technology has “a valuable history of putting things at our fingertips to be seized and used widely for their designed objectives as well as for other purposes” (p. 11). However, he warned that the opportunities provided by computer technology are not always accepted in education. He further explained that computer technology actually has “two order fingertip effects.” The first order fingertip effects occur when a computer innovation changes “the way people do certain things without actually changing very much the basic aspirations, endeavors, or thinking habits

“At a time when traditional assessment is under increasing scrutiny and criticism, the nation is placing greater expectations on the potential role of the computer in educational assessment.”

***“Computer technology
does a great job
of automating
varying phases of
testing processes
such as creating,
storing, distributing
and sharing test
materials.”***

of a population” (Perkins, 1985, p. 11). For instance, unlike regular mail, emails and instant messaging allow for faster communication with friends, relatives and business associates thousands of miles away. Another example is that computer-based tests use built-in databases to automatically collect and compute data. In these instances, the first order fingertip effects of computer technology answer the question, “What could you do that you could not before?” Specifically, computer technology

can help automate routine procedures quickly and accurately, thus improving productivity and efficiency.

The second order fingertip effects answer the question, “What difference will a computer *really* make to a person’s higher-order skills, i.e., decision making, reflection, reasoning and problem solving?” (Perkins, 1986, p. 11). Jonassen (2000) indicated that the second order effects should help “in the construction of generalizable, transferable skills that can facilitate

thinking in various fields” (p. 18). However, when computer technology is used, it should go beyond its automated function as a production tool; it should be used to promote higher-order skills. For example, Microsoft Excel is a spreadsheet tool that is useful for teachers as a grade book. By inputting grades and one function, or a series of functions, a teacher can produce report cards very quickly, thus saving time at the end of a marking period. The spreadsheet data can also be turned into a graph so that the report card can be seen visually or graphically. The teacher can also look at the results and use them to work on comments for student performance. In doing so, the teacher can reflect on his or her teaching, ask “what if” questions or help a struggling student. In this instance, having the computer do the menial task of averaging grades and displaying graphics are the first order fingertip effects; reflective thinking and helping a struggling student’s learning are the second order.

In summarizing Perkins’ words, if computer technology fails to achieve its full potential, it has only been used to achieve the first order fingertip effects. It is the second order fingertip effects, the non-automatic and effortful process, that establishes the true value of the computer technology. With this distinction in mind, the

question then becomes: Which level of fingertip effects of the computer have current computer-based assessment tools achieved?

***Limitations of computer technology
use in assessment***

Computer applications in educational assessment are evident in testing preparation, administering, scoring and reporting (Zenisky & Sireci, 2002). To this end, computer technology is often used to present test items and collect responses. Clearly, computer technology does a great job of automating varying phases of testing processes such as creating, storing, distributing and sharing test materials. The automation, especially in large-scale assessment such as that administered by the Educational Testing Service (ETS), can benefit both examiners and examinees in multiple ways.

Unlike most computer-assisted tests, Computer Adaptive Testing (CAT), which has been used and improved during the past 15 years, has noteworthy advantages over “fixed-item” tests. This adaptive approach to testing can update the estimate of an examinee’s ability after each test item and select the appropriate level of subsequent items for the examinee. In this way, student deficiencies and strengths can be quickly identified and addressed. Another innovative use of computer-based assessment can be seen in some computer simulation projects. In their project, Shavelson, Baxter and Pine (1992) required their students to replicate electric circuits by manipulating icons of batteries and wires presented on a Macintosh computer.

Obviously, CAT and computer simulations demonstrate a more sophisticated approach to testing, but these strategies are seldom implemented with teacher-made tests due to technical complexity and logistical problems. Instead, teachers often use computers to help such small-scale assessments as creating traditional multiple-choice, fill in the blank and short essay type questions. It is worth pointing out that in either large or small-scale assessments, computer technology is frequently used as a test preparation and production tool (Perkin’s first order fingertip effects) rather than as a learning tool to enhance higher-order skills (the second order fingertip effects).

Another observation of computer-based assessment is that computer technology is often seen by teachers as a “representation container” rather than an effective assessment tool (Dede, 2003, p. 7). Specifically, computers allow students to create multimedia materials at any time and on demand and, as a result, make learning visual, mobile and fun. However, after

students finish the projects, it is not unusual for students' competencies in using computer software to be measured while the generic problem-solving competencies are ignored (Becker & Lovitts, 2003, p. 134). Obviously, it is much easier to assess computer literacy than the problem-solving process. However, when assessment strategies only play a secondary role and the real outcomes of the projects' generic problem-solving are not assessed, chances are that computers are not being used to their full potential by teachers.

The third observation is that the design of the computer-based assessment does not "adequately support human practices that produce meaningful information about student learning" (Hall, Knudsen, & Greeno, 1996, p. 316). Take the design of multiple-choice tests for example. Multiple-choice questions are the most widely used format in computer-based assessments, but this format has been criticized for giving students no practice at expressing their thoughts and for not providing individual feedback or interactions regarding student performance. It is important to note that assessment approaches could easily replace one form of computer technology with another without really paying attention to human interactions. For example, portfolios are used in place of standardized examinations, whereas little explicit attention is paid to the human interactions surrounding either the portfolio or the standardized examinations. In fact, an effective assessment requires extensive interactions between examiners and examinees (Bahr & Bahr, 1997). Such interactions provide an opportunity for examiners to identify learning gaps and for examinees to moderate their learning.

Emerging technology assessment tools

A review of the related computer literature would indicate that the current use of computer technology in educational assessment does not achieve Perkins' (1985) second order fingertip effects in general. As discussed above, the second order should go beyond the automated function of computer technology and extend to enhance higher-order skills. Fortunately, although still in their early stage of development, some emerging prototype tools have demonstrated great potential to push computer-based assessment beyond the automation of testing, representation channels and insufficient human practices.

The SMART model provides one example. This computer-based learning tool for science and math concepts contains a variety of assessment strategies permeating a problem-based

and project-based learning environment. On-going assessment is incorporated throughout the learning process in a way that allows computer technology to support student reflection. In addition, student works are evaluated by self-assessment, peers, teachers and external agencies. This way teachers can identify deficiencies and strengths in student performance. Equally important, students can reflect on their learning process and improve their higher-level skills. Table 1 reviews a selection of other prototype tools in which assessment strategies are interwoven with the learning process.

Discussion and conclusions

Computer technology has revolutionized instruction and student learning, and it holds great promise for enhancing educational assessment. Although still in their early stages of development, computer-based assessment tools offer innovative approaches for documenting students' learning process, identifying learners' deficiencies and strengths, and providing timely feedback. Such a promise cannot be realized without the cooperation of instructional technologists, teachers and schools.

It is true that computer technology is as powerful as it is seductive. It is easy for instructional technologists to get carried away and spend all their time designing scenarios and gathering complex data, only then to ask "How do we assess it?"

(Mislevy, Steinberg, Almond, Haertel, & Penuel, 2003). When this happens, computer technology is not used to its maximum potential. With rapid advances in computer technology, the challenge for instructional technologists is to capture more complex performances in assessment settings. To design effective complex assessments, instructional technologists should read Messick's (1994) discussion about computer-based simulations, portfolio assessments and performance tasks.

It is also true that it is beyond many teachers' abilities to design advanced assessment prototype tools. In fact, the concept of Perkins' second order fingertip effects can be applied to many classroom assessment routines, especially with the help of free internet resources. For example, when incorporating the idea of

"With rapid advances in computer technology, the challenge for instructional technologists is to capture more complex performances in assessment settings."

| Tools | Area | Assessment Strategies | Source |
|----------------|---|---|-----------------------------|
| DIAGNOSER | Key math and science concepts | <p>This web-based program contains the following tools and strategies to assess student learning and provide feedback:</p> <p>DIAGNOSER: Students receive ongoing feedback as they work through their assignment. Teachers receive a summary of student diagnoses.</p> <p>Elicitation Questions: After students respond to the carefully constructed questions, the program can pinpoint areas of possible misunderstanding, give immediate feedback on reasoning strategies, and prescribe relevant instruction.</p> <p>Developmental Lessons: These lessons open up the ideas elicited in the class discussion and help students test their initial ideas.</p> <p>Prescriptive Activities: Teachers can use activities to target specific problematic ideas.</p> | DIAGNOSER Tools Website |
| eduPortfolios | Digital portfolio | <p>This tool allows intimate interaction between students, teachers and other stakeholders. On the one hand, students can view and assess real student work and compare them against established learning standards. On the other hand, students are asked to write about how they understand the learning standards and how they meet the standards in their work.</p> <p>Afterwards, feedback from multiple teachers is attached to student portfolios and their reflections. In this way, students can see how their understanding matched or did not match their teachers' understandings, and vice versa. This approach allows for a continual process of reflection, understanding and learning.</p> | Ahn, 2004 |
| Summary Street | Reading comprehension and writing skills | <p>Summary Street is an educational software based on latent semantic analysis (LSA), which is a computer method for representing the content of texts.</p> <p>Students can prepare multiple drafts of a summary and receive content-based feedback. For example, <i>the Redundancy Check</i> performs a sentence-by-sentence comparison to flag sentences that appear to have overlapping content. <i>The Relevance Check</i> compares each sentence with the original text and pinpoints sentences that have little or no relevance to the topic.</p> <p>In addition, the content feedback is presented in a game-like, easy-to-grasp graphic display. In this way, students are more willing to repeat the cycles of rewriting and revision before submitting the final summaries to their teachers.</p> | Wade-Stein & Kintsch (2004) |
| PROUST | Programming problems in using the Pascal language | <p>PROUST is considered a milestone in the field of intelligent tutoring systems. This system describes a diverse array of programming problems and the ways in which parts of the problem can be solved.</p> <p>Based on how people reason out computer programs, PROUST is designed to analyze a student's computer program, identify strengths and weaknesses in the student's work, and then present comments on the student's work. This software not only can identify errors of syntax, but, more interestingly, on errors in a student's solutions to solve a problem.</p> | Johnson & Soloway (1985) |

Table 1. Prototype tools of computer-based assessment in education.

computer-assisted tests into existing teacher practice, teachers can search for creative and effective methods for conducting testing and evaluation in addition to the traditional multiple choice, fill in the blank and short essay questions (Khan, 1997). They can include web-based group discussions and e-portfolio development to evaluate students' progress. They can allow students to submit comments and reflections about their project design and delivery activities via a web log (or blog, a type of online learning diary). They can also use computer simulations for hands-on performance assessments. All these assessment strategies can be greatly facilitated by using free online resources.

Another example of applying Perkins' second order fingertip effects is the use of online rubric tools such as Rubistar. In fact, teachers can relinquish their intellectual authority a little and have students create the rubric in groups. By negotiating the rubric among their peers and with their teacher, students can spell out their project expectations and have ownership of the assessment process. Afterwards, the students can use the rubric as a central guidance to provide feedback to their counterparts. In this instance, the use of rubric tools for communication, negotiation and peer review is in line with Perkins' second order fingertip effects. It is toward this end that the teachers and students can make the most of computer-based tools for assessment.

Lastly, what are the roles of schools in incorporating Perkins' second order fingertip effects into computer-based assessment practices? Some may argue that many schools do not have the technology infrastructure and/or the budgets to support the effort of incorporating computer-based tools into classroom assessments. The real challenge, however, is to overcome the fear, suspicion and doubt that are found in many schools about the relative importance of such efforts. The point cannot be made more clearly than

Dede (2003) did when he claimed that "the fundamental barriers to employing these technologies effectively for learning are not technical or economic, but psychological, organizational, political and cultural" (p. 9).

Hong Lin received her doctoral degree from the Department of Learning and Performance Systems at Penn State, University Park campus. She is manager of faculty development in the Institute for Teaching and Learning Excellence at Oklahoma State University. Her research interests include, but are not limited to, online instruction, blended learning, assessment and the ethical applications of instructional technology in higher education.

Dr. Frank Dwyer is professor of education in the instructional systems program in the Department of Learning and Performance Systems at Penn State. Dr. Dwyer was president of the Association for Educational Communications and Technology (AECT) from 1984 to 1985. He led the Department of Adult Education, Instructional Systems and Workforce Education and Development at Penn State from 1990 to 1995. Dr. Dwyer's research interests focus on distance education, corporate instructional systems, instructional design/strategies and visual learning systems.

References

- Ahn, J. (2004, April). Electronic portfolios: Blending technology, accountability and assessment. *T.H.E. Journal*, 31(9), 12-18. Retrieved April 20, 2005, from <http://thejournal.com/magazine/vault/A4757B.cfm>
- Baek, S. G. (1994). Implications of cognitive psychology for educational testing. *Educational Psychology Review*, 6(4), 373-389.
- Bahr, M. W., & Bahr, C. M. (1997, Winter). Educational assessment in the next millennium: Contributions of technology. *Preventing School Failure*, 41(2), 90-94.
- Becker, H. J., & Lovitts, B. E. (2003). A project-based approach to assessing technology. In G. D. Haertel & B. Means (Eds.), *Evaluating educational technology* (pp. 129-148). New York, NY: Teachers College Press.
- DIAGNOSER TOOLS. Retrieved September 1, 2005, from <http://www.diagnoser.com/diagnoser/>
- Dede, C. (2003, March-April). No cliché left behind: Why education policy is not like the movies. *Educational Technology*, 43(2), 5-10.
- Hall, R. P., Knudsen, J., & Greeno, J. G. (1996). A case study of systemic aspects of assessment technologies. *Educational Assessment*, 3(4), 315-361.
- Johnson, W. L., & Soloway, E. (1985). PROUST: An automatic debugger for PASCAL programs. In P. Lemmons (Ed.), *Lecture notes in computer science* (pp. 179 - 190). Hightstown, NJ: McGraw-Hill, Inc.
- Jonassen, D., H. (2000). *Computers as mindtools in schools: Engaging critical thinking* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall, Pearson Education.
- Khan, B. H. (1997). Web-based instruction: what is it and why is it? In B. H. Khan (Ed.), *Web-based Instruction* (pp. 5-18). Englewood Cliffs, NJ: Educational Technology.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23(2), 13-23.
- Mislevy, R. J., Steinberg, L., S., Almond, R., G., Haertel, G., D., & Penuel, W. R. (2003). *PADI technical report 2: Leverage points for improving educational assessment*. SRI International.
- Perkins, D. (1985, August/September). The fingertip effect: How information-processing technology shapes thinking. *Educational Researcher*, 14, 11-17.
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Research Council.
- Shavelson, R. J., Baxter, G. P., & Pine, J. (1992). Performance assessments: Political rhetoric and measurement reality. *Educational Research* (21), 4, 168-177.
- Wade-Stein, D., & Kintsch, E. (2004). Summary Street: Interactive computer support for writing. *Cognition and Instruction*, 22(3), 333-362.
- Zenisky, A. L., & Sireci, S. G. (2002). Technological innovations in large-scale assessment. *Applied Measurement in Education*, 15(4), 337-362.